Section 1-4 and 1-5: Scatterplots, Lines of Best Fit, and Correlation

By the end of this lesson, you should be able to answer:

- How do you use scatterplots to solve problems?
- How do you use a graphing utility to determine a line of best fit?

Where you might see this in the real world:

- Retail, education, sports, statistics, insurance

Define the following words:

1. Scatterplot

2. Positive Correlation

3. Negative Correlation

4. Line of Best Fit

5. Trend Line

We have all seen scatterplots. They are basically just collections of points on the coordinate plane. But how do we graph them on the calculator?

The following information shows the amount of time spent exercising each week and the resting heart rates in beats per minute of members of an aerobics class.

4 h, 60 bpm	3 h, 63 bpm	3.5 h, 67 bpm	4 h, 55 bpm
2 h, 70 bpm	2.5 h, 65 bpm	3 h, 60 bpm	5 h, 50 bpm
4.5 h, 60 bpm	5 h, 65 bpm	4.5 h, 50 bpm	3 h, 70 bpm
5 h, 55 bpm	1 h, 75 bpm	3 h, 55 bpm	4 h, 57 bpm

This information can be collected into a set of ordered pairs, where the first value will be time in hours, and the second value will be beats per minute. Collect the above data into a set of ordered pairs.

(4, 60),

Now we want to enter this information into our calculators. We will enter the values for time in **L1**, and bpm in **L2**. Make sure that you enter the values in the same order for both lists (that is, make sure that the points "stay together."

Now press **2nd** then **STAT PLOT** to graph this information. Go to **1:Plot1…** and press **ENTER**. Turn this plot **On** and select the scatterplot under **Type:**. Your **Xlist:** should be **L1** and **Ylist:** should be **L2** (or whatever lists you may have used). You can also choose which mark will appear on your graph. Next, press **ZOOM** and find **9:ZoomStat**. You now have your scatterplot shown, without having to worry about setting up a window. **NOTE:** You may not get the graph you want if you have an equation listed under **Y=**, so make sure all of those graphs are turned off.

You will notice there is a negative correlation on this scatterplot, though it is not a strong one. We can check this by making a line of best fit. First, we need to turn on the diagnostic. Press **2nd** then **CATALOG** and find **DiagnosticOn** and press **ENTER.** Now press **STAT** and go to the **CALC** menu. Find **4:LinReg (ax+b)** and press **ENTER**. We need it to find the line of best fit for the lists **L1** and **L2**, so we need to enter those on our screen as well, placing commas after each list. We also need to have the equation, so we press **VARS** and go to the **Y-VARS** menu and find **1:Function…**. We want **1:Y1.** Now your screen should have this on it: **LinReg(ax+b) L1, L2, Y1**. Press **ENTER**. You will see on the main screen what y, a, b, r^2, and r all equal. The correlation will be given by r. The closer to 1 or −1, the better the correlation.

Now, if you press **Y=**, you will see an equation for the line of best fit. Also, if you press **GRAPH**, you will see the line on our scatterplot. We can now use this line to predict what will happen at other points on the graph.

Predict the heart rate is for someone who exercises for 2 hours:

Example 1: The following data represents the average time studying per week in hours with the final grade earned in a math class. Create a scatterplot and put in a line of best fit, then answer the questions.
(6, 80), (1, 65), (5, 85), (2, 60), (5, 70), (3, 70), (9, 95), (8, 90), (3, 80), (7, 70), (4, 65), (8, 75), (4, 80), (7, 85), (9, 85), (10, 95), (1, 75)

a. What correlation exists?

b. What is the correlation value (r)?

c. If you spent 5 hours a week studying, what grade would you expect?

Question: How strong should a correlation be in order to show a strong relationship?

Problem Set: